

## Forschungsdaten

AR 3247

---

*Annette Strauch und Volker Hess*

# Von der Produktion bis zur Langzeitarchivierung qualitativer Forschungsdaten im SFB 1187

**Zusammenfassung:** Im Rahmen des DFG-SFB 1187, „Medien der Kooperation“, an der Universität Siegen, in dem kooperative Praktiken im Mittelpunkt stehen, beschäftigt sich das Zentrum für Informations- und Medientechnologie (ZIMT) zusammen mit der Universitätsbibliothek mit forschungsbegleitenden Hilfestellungen für die einzelnen Forscher in den Teilprojekten beim Umgang mit ihren „Forschungsdaten – von der Produktion bis zur Langzeitarchivierung“. Das Zentrum für Informations- und Medientechnologie (ZIMT) unterstützt seit Jahren insbesondere die medienwissenschaftliche Forschung an der Universität Siegen. Die vielfältigen Untersuchungen von kooperativen Medien im SFB 1187 führen zu einer Sammlung großer und heterogener Datenmengen über hoch kontextbezogene Kooperationszusammenhänge, die es zu verwalten gilt, welches eine große Herausforderung für ein nachhaltiges Forschungsdatenmanagement darstellt. Für die Forscher im SFB müssen unterstützende Forschungsdateninfrastrukturen konzipiert werden. Auf dem Poster werden verschiedene fachspezifische und disziplinübergreifende Anforderungen an das Forschungsdatenmanagement im DFG-SFB 1187 gezeigt (Daten, Datenanalyse, Tools, Speicherung, etc.) und analysiert. Dargestellt werden unterschiedliche Use-Cases.<sup>1</sup>

**Schlüsselwörter (dt.):** Forschungsdaten, Forschungsdatenmanagement, DFG-Sonderforschungsbereich, INF-Projekt, qualitative Forschungsmethoden, Sozial- und Geisteswissenschaft, Medienethnologie

## **From production to long-term archiving of qualitative research data in the CRC 1187**

**Abstract:** As part of the German Research Foundation (DFG) – Collaborative Research Center 1187, “Media of Cooperation” at the University of Siegen in which cooperative practices are central, the Center for Information and Media Technology (ZIMT) focuses on the research supporting assistance for the individual researchers in the sub-projects dealing with their “research data - from production to long-term archiving”. The Center for Information and Media Technology (ZIMT) has been

---

<sup>1</sup> Dieser Text ist eine Überarbeitung des Posterbeitrags von Annette Strauch und Volker Hess auf der 18. DINI-Jahrestagung „Forschungsdaten – von der Produktion zur Langzeitarchivierung“ am 04. und 05. Oktober 2017 in der Niedersächsischen Staats- und Universitätsbibliothek Göttingen: <https://dini.de/veranstaltungen/jahrestagungen/2017>.

supporting media research in particular at the university for years. The multifaceted examination of cooperative media in the CRC 1187 lead to a collection of large and heterogeneous data sets in highly contextual collaborative relationships that need to be managed, which is a major challenge for a sustainable research data management. For the researchers at the CRC Researchers supportive research data infrastructures are needed to be developed. The poster shows various subject-specific and interdisciplinary requirements for research data management in the DFG-CRC 1187 (data, data analysis, tools, storage, etc.) and analyses them.

**Keywords:** Research Data, Research Data Management, DFG-Collaborative Research Center, INF-project, qualitative research methods, Social Sciences and Humanities, Media Ethnology

## **1 SFB 1187 „Medien der Kooperation“**

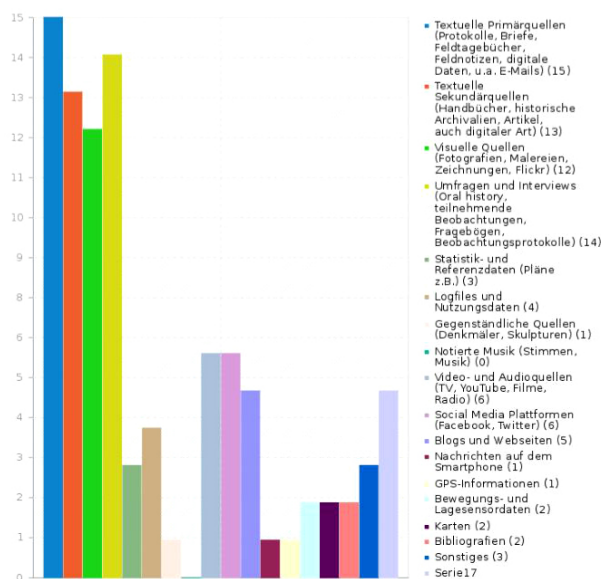
Seit Mai 2016 leisten Forscher aus den Geistes-, Kultur-, Sozial- und Ingenieurwissenschaften einen Beitrag zum Verständnis der digitalen Gegenwartskultur im Sonderforschungsbereich „Medien der Kooperation“ (SFB 1187)<sup>2</sup> der DFG. Auf dem Poster werden unterschiedliche Use-Cases mit jeweiligen Workflows in Anlehnung an das Data Continuum aus Ergebnissen einer Umfrage und aus Interviews dargestellt. Bedarfe für die Wissenschaftler aus den Geistes-, Kultur-, Sozial- und Ingenieurwissenschaften sowie der Siegener Sozioinformatik sind untersucht worden. Das ZIMT greift bei seiner Arbeit zum einen auf Erfahrungen im Forschungsdatenmanagement zurück (DINI, DARIAH, RADAR, RDA, nestor, Projekte), zum anderen sollen fachspezifische Anforderungen in Kooperation von ZIMT und Universitätsbibliothek Siegen in zentrale Serviceleistungen (technische Services vom ZIMT) umgewandelt werden, so dass diese auf die Forschergruppen im DFG- SFB 1187, und darüber hinaus, entsprechend zugeschnitten sind. Die Entwicklung von Verfahren, Standards, technischen Infrastrukturen und Werkzeugen soll also projektabhängig und interdisziplinär entwickelt werden, um die bestmöglichen Voraussetzungen sowohl für ein sinnvolles Forschungsdatenmanagement eines einzelnen Produzenten im SFB als auch für kollaborative Forschungsk Kooperationen bereitzustellen.

Generell ist der Charakter der Forschungsdaten qualitativer Art. Es handelt sich hier hauptsächlich um textuelle Primärdaten (Protokolle, Briefe, Feldtagebücher, Feldnotizen, digitale daten, E-Mails), textuelle Sekundärquellen (Handbücher, historische Archivalien, Artikel), Visuelle Quellen (Fotografien, Malereien, zeichnungen, Flickr), Umfragen und Interviews (Oral history, teilnehmende Beobachtungen, Fragebögen, Beobachtungsprotokolle) und vor allem häufig um Social Media Daten.

---

<sup>2</sup> <https://www.mediacoop.uni-siegen.de/de>.

Aus welchen Quellen schöpfen Sie ihre Forschungsdaten hauptsächlich?



**Abb. 1:** Datenquellen des SFB 1187, LimeSurvey Frühjahr 2017

Beteiligte Einrichtungen des SFB 1187 sind neben der Universität Siegen (alle Fakultäten), auch die Universität zu Köln (a.r.t.e.s Graduate School for the Humanities), das Jung-Stilling-Krankenhaus, Siegen, das Siegerlandmuseum und das Museum für Gegenwartskunst in Siegen.<sup>3</sup>

## 2 Herausforderung durch Heterogenität: Hintergründe, Aufgaben und Potentiale des Projekts SFB 1187-INF

Das INF-Modul „Infrastrukturkonzepte für die Beforschung kooperativer Medien“<sup>4</sup> hilft u.a. beim Aufbau technischer Infra- und Beratungsstrukturen zum FDM, die vom Zentrum für Informations- und Medientechnologie (ZIMT) und der Universitätsbibliothek der Universität Siegen getragen werden und sich dabei an der universitären Forschungsdaten-Policy orientiert.<sup>5</sup> Die sozio-technische Untersuchung und Gestaltung kooperativer Forschungspraktiken im SFB ist ein weiterer zentraler Aspekt von INF. Hier wird die interdisziplinäre und kooperative Forschungspraxis in einer akteur- und aktivitätsorientierten Weise dokumentiert. Kooperative Forschungsmedien werden im Gestaltungslabor behandelt und Forschungswerkzeuge werden untersucht. Eine Austauschplattform für neue Methoden der Datenerhebung und -analyse wurde bereitgestellt.

3 a) beteiligte Einrichtungen: Jung-Stilling-Krankenhaus, Siegerlandmuseum und Museum für Gegenwartskunst in Siegen sind beteiligte Einrichtungen; b) Standorte: Universität zu Köln (a.r.t.e.s Graduate School für the Humanities) und Universität Mannheim sind Standorte des SFB, also offizielle Kooperationspartner mit Stellenanteilen.

4 <https://www.mediacoop.uni-siegen.de/de/projekte/inf>.

5 Forschungsdaten-Policy der Universität Siegen: [https://www.universi.uni-siegen.de/usi\\_forschungsdaten-policy-1.pdf](https://www.universi.uni-siegen.de/usi_forschungsdaten-policy-1.pdf).

Die methodische Vielfalt ist neben der Unterschiedlichkeit der Forschungsdaten der im Sonderforschungsbereich 1187 „Medien der Kooperation“ versammelten Teilprojekte<sup>6</sup> als eine große Herausforderung für das Datenmanagement qualitativ ausgerichteter geistes-, kultur- und sozialwissenschaftlicher Forschung hervorzuheben. Orientierungsgespräche im Rahmen von „Gestaltungslaboren“ („Research Tech Labs“) als offene Foren zur Diskussion erforderlicher methodischer und infrastruktureller Lösungsansätze zeigten, dass ohne die Verständigung über die Notwendigkeit eines strukturierten Umgangs gerade mit sensiblen, insbesondere personenbezogenen Daten der spezifischen Forschungsrichtung, und eine systematische Anforderungserhebung das Ziel, generische „Basisservices zur Datenspeicherung und zum Datenmanagement“ zu adaptieren und zu entwickeln, nicht umzusetzen sein würde. Es fand zunächst ein Dialog mit allen Akteuren der insgesamt 13 Teilprojekte des SFB 1187 statt. Die Klärung der Bedarfe orientiert sich bei diesem Austausch eng an den Forschenden sowohl in geschlossener Umfrage (LimeSurvey) wie auch in teilstandardisierten Interviews. Hierbei entstanden generalisierbare Anforderungen und Lösungen sowie projektspezifische Anforderungen und Lösungen. Die Anforderungen werden immer aus den „Communities of Practice“, die sich an den Forschern orientieren, heraus betrachtet, wobei das „Curation Domain Model“ zur Beschreibung des Lebenszyklus von Forschungsdaten nach Treloar et al.<sup>7</sup> für einzelne, möglichst exemplarische Anwendungsfälle als Orientierungsrahmen im Forschungsdatenmanagement herangezogen wird. Innerhalb von INF wurde außerdem untersucht, wie Forschende Online-Profile (Webseiten, ResearchGate, Academia.edu, Google Scholar, etc.) nutzen, um sich (a) selbst darzustellen und (b) über andere Forschende zu informieren? Welche Informationen werden herangezogen, um sich ein Bild über Epistemologien, Theoriebezüge und Methoden der Forschenden zu machen? Ein weiterer Fokus der empirischen Arbeiten sind Motivationen und Nutzungsgeschichte von Forschungswerkzeugen.

### **3 Untersuchung der Nutzercommunities im INF-Projekt**

An deutschen Hochschulen sind die von der DFG geförderten INF-Projekte<sup>8</sup> zu sehr wichtigen Impulsgebern im Forschungsdatenmanagement geworden.<sup>9</sup> Die Forschenden im SFB 1187 wurden zunächst durch Veranstaltungen wie dem Gestaltungslabor<sup>10</sup> mit den Akteuren im Forschungsdatenmanagement, den Forschungsforen und Jahrestagungen, zum

---

6 <https://www.mediacoop.uni-siegen.de/de/projekte>.

7 Treloar et al. (2007).

8 DFG-Vordruck 50.06.-07/18 Merkblatt Sonderforschungsbereiche, zuletzt geprüft am 25.08.2018, [http://www.dfg.de/formulare/50\\_06/50\\_06\\_de.pdf](http://www.dfg.de/formulare/50_06/50_06_de.pdf)

9 DFG-Vordruck 50.06.-07/18 Merkblatt Sonderforschungsbereiche: [http://www.dfg.de/formulare/50\\_06/50\\_06\\_de.pdf](http://www.dfg.de/formulare/50_06/50_06_de.pdf).

10 <https://www.mediacoop.uni-siegen.de/de/projekte/inf/#gl>.

Forschungsdatenmanagement sensibilisiert. Anhand der Freitextfelder der unveröffentlichen LimeSurvey zum Forschungsdatenmanagement im Frühjahr 2017 für die einzelnen Teilprojekte im SFB 1187 war zunächst zu erkennen, dass der Nutzen vom Datenmanagement, der Mehrwert dessen, oft noch nicht von den Forschenden erkannt wird.

„Da ich Ethnologin bin, ist Forschungsdatenmanagement sicherlich weniger zentral.“

“Ich kann überhaupt nicht erkennen, wie andere ForscherInnen von meinen Daten profitieren könnten. Umgekehrt fehlt mir die Phantasie mir vorzustellen, wie ich von Daten anderer Forscher profitieren sollte.”

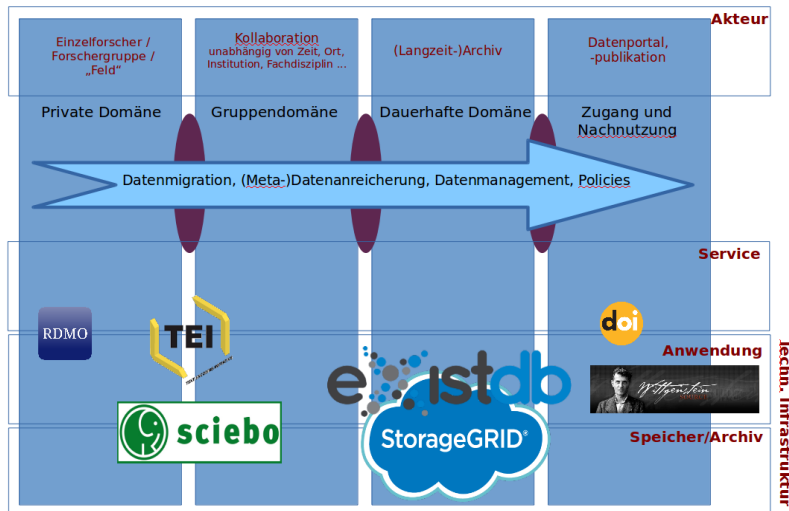
#### **4 Generische Infrastrukturen und projektspezifische Infrastrukturen**

Dem ZIMT als universitärem Infrastrukturpartner des Sonderforschungsbereichs stellen sich die Herausforderungen zum einen nachhaltige generische und zum anderen nachhaltige projektspezifische Infrastrukturen anzubieten, die in Kooperation mit der Universitätsbibliothek und anderen Akteuren an der Universität Siegen gestaltet und im Rahmen des Teilprojekts INF den Teilprojekten vermittelt werden. Die technischen Infrastrukturen orientieren sich an einem offenen Architekturmodell mit möglichst standardisierten Schnittstellen. Ziel ist die flexible Integration diverser, möglichst generischer Komponenten zur Unterstützung der Forschenden bei der Umsetzung sehr spezifischer Praktiken der Datenerhebung, -aufbereitung, -analyse, -pflege und -publikation. Zugleich sind Werkzeuge implementiert, die die Planung von Datenmanagementprozessen im Rahmen dieser modularen technischen Landschaft unter Berücksichtigung und Beteiligung weiterer institutioneller Partner (u.a. Datenschutz, wissenschaftliche Weiterbildung, Forschungsförderung etc.) möglichst frühzeitig im Forschungsprozess ermöglichen. Übergreifende Handlungserfordernisse konnten in allen Teilprojekten ermittelt werden: Lösungen für Kollaboratives Arbeiten<sup>11</sup> an auf nachhaltigen, sicheren Arbeitsplattformen verfügbar gehaltenen Forschungsdaten müssen weiter entwickelt werden, die Sicherstellung von deren kontrollierter, insbes. rechtssicherer Nachnutzung im Projekt muss gewährleistet sein ebenso wie in diversen übergreifenden Archivierungs- und Publikationskontexten. Teilweise große Unsicherheiten bzgl. der rechtlichen, organisatorischen und technischen Randbedingungen dieser forschungsbegleitenden Verfahren wurden innerhalb der einzelnen Teilprojekte formuliert und offenbarten einen hohen Beratungsbedarf bei den handelnden Akteuren. Generische technische Infrastrukturen verorten sich bislang vorrangig in der Gruppen- und Dauerhaften Domäne gemäß dem „Curation Domain Model“ (s.o.). Dazu gehören Werkzeuge zur begleitenden Metadatenerfassung und -anreicherung in den genannten Arbeitsumgebungen und dem einfachen Ingest in Nachweissysteme, Langzeitarchive und Publikationsplattformen, die

---

11 <https://www.forschungsdaten.info/themen/organisieren-und-arbeiten/kollaboratives-arbeiten>.

möglichst an realer wissenschaftlicher Praxis erprobt intuitive Workflows ermöglichen. Am Beispiel des Teilprojekts „Wissenschaftliche Medien der Praxistheorie: Harold Garfinkel und Ludwig Wittgenstein“ (P01)<sup>12</sup> lässt sich eine projektspezifische Erweiterung der generischen Infrastrukturen in Orientierung am Rahmenmodell unter Einbindung externer Komponenten (z.B. TEI) und Dienstleister (Softwarefirmen) sowie wissenschaftlicher Partnerinstitutionen verdeutlichen.



Use Case: Digitale Editionen (exemplarisch)

Basis: Treloar, Domänenmodell; FoDaKo-Poster DINI JT 2017

**Abb. 3:** „Curation Domain Model“ für den exemplarischen Use-Case: Digitale Edition Wittgenstein

## 5 Von der Produktion bis zur Langzeitarchivierung – Ausblick

Mit allen Projekten des SFB 1187 sind von Beginn des INF-Projekts an gute Kontakte zum Forschungsdatenmanagement zustande gekommen. Die Forscher haben sich mit dem Thema Forschungsdatenmanagement auseinandergesetzt. INF, das ZIMT sowie die Universitätsbibliotheken haben gelernt, worin die Anforderungen und Bedarfe nach Services und Infrastrukturen liegen. Sie müssen ihre Services hier weiter ausbauen, um den Sonderforschungsbereich im Sinne der guten wissenschaftlichen Praxis bestmöglich zu unterstützen, um damit eine exzellente Forschung und die Bewahrung von deren Erträgen zu gewährleisten. Die ermittelten Herausforderungen für Datenmanagement, -archivierung und -nachnutzung im spezifischen Forschungsfeld des SFB sind bislang nur im Ansatz systematisch beschrieben, Lösungsansätze bilden sich erst langsam als „Best Practices“ heraus und erfordern für die nachhaltige Umsetzung fach- und einrichtungsübergreifende Abstimmung und Kooperation.

Erfahrungen, Herausforderungen und Spannungsfelder aus dem SFB 1187 beim Umgang mit qualitativen Daten wurden beim RatSWD Workshop „Archivierung und Zugang zu qualitativen

<sup>12</sup> <https://www.mediacoop.uni-siegen.de/de/projekte/p01>.

Daten“, der im Frühjahr 2018 in Bremen stattfand, diskutiert.<sup>13</sup> Benannt wurden bei dem Bremer Workshop diverse Anforderungen für eine bessere Archivierungsinfrastruktur in der qualitativen Forschung. Diskutiert wurde über den Aufbau und der Nutzung professioneller Infrastrukturen und Forschungsdatenzentren. Die Beiträge der gemeinsam erarbeiteten Empfehlungen sollen Ende 2018 veröffentlicht werden als weiterer Schritt bei der Lösung für eine institutionelle Struktur für die Verfügbarmachung und die Archivierung qualitativer Forschungsdaten. Mit dem Archiv für qualitative Forschungsdaten bzw. mit dem dafür geschaffenen nationalen Datenzentrum (DSZ), das am Socium-Forschungszentrum Ungleichheit und Sozialpolitik angesiedelt ist, wird die Forschungsdateninfrastruktur für die empirische Sozialforschung nachhaltig verbessert. Die Deutsche Forschungsgemeinschaft hat im April 2018 die Implementation von QUALISERVICE<sup>14</sup> als Datenservicezentrum (DSZ) für qualitative sozialwissenschaftliche Forschungsdaten bewilligt. So wurde der Weg geebnet, dass Sozialforscher in Deutschland erstmals eine verlässliche Anlaufstelle erhalten, um Interviewdaten archivieren und für die Sekundärnutzung in verlässliche Anlaufstelle erhalten, um Interviewdaten archivieren und für die Sekundärnutzung in Forschung und Lehre anfordern zu können. In den oben erwähnten Interviews im Zusammenhang mit der Beforschung kooperativer Medien und Forschungsarbeiten aller SFB Teilprojekte durch INF, wurde mit Forschenden aus einzelnen Projekten über die Möglichkeiten und Grenzen eines nachhaltigen Forschungsdatenmanagements im Zusammenhang mit Qualiservice anhand aller Phasen im Datenzyklus diskutiert. Berücksichtigt wurden hierbei z.B. Aspekte des Datenschutzes, möglicher Veröffentlichungen von Daten als eigenständige Publikationen sowie Zugänge zu Forschungsdaten aus dem SFB unter Berücksichtigung verschiedener Nutzungsszenarien (Closed vs. Open z.B).

### **Literaturverzeichnis**

Treloar, A.; Groenewegen, D.; Harboe-Ree, C. (2007): The Data Curation Continuum. In: *D-Lib Magazine*, 13(9/10), 1082–9873.

---

<sup>13</sup> <https://www.ratswd.de/events/QualiWS2018>.

<sup>14</sup> <http://www.qualiservice.org>.



**Annette Strauch**

Forschungsdatenmanagement

Universitätsbibliothek Hildesheim an der Stiftung Universität Hildesheim

Universitätsplatz 1

D-31141 Hildesheim

[annette.strauch@uni-hildesheim.de](mailto:annette.strauch@uni-hildesheim.de)



**Volker Hess**

Zentrum für Informations-und Medientechnologie ZIMT an der Universität Siegen

Hölderlinstraße 3

D-57076 Siegen

[volker.hess@uni-siegen.de](mailto:volker.hess@uni-siegen.de)